

Cooperative Change Detection for Online Power Quality Monitoring

Shang Li[†] and Xiaodong Wang[†]

Abstract

This paper considers the real-time power quality monitoring in power grid systems. The goal is to detect the occurrence of disturbances in the nominal sinusoidal voltage/current signal as quickly as possible such that protection measures can be taken in time. Based on an autoregressive (AR) model for the disturbance, we propose a generalized local likelihood ratio (GLLR) detector which processes meter readings sequentially and alarms as soon as the test statistic exceeds a prescribed threshold. The proposed detector not only reacts to a wide range of disturbances, but also achieves lower detection delay compared to the conventional block processing method. Then we further propose to deploy multiple meters to monitor the power signal cooperatively. The distributed meters communicate wirelessly to a central meter, where the data fusion and detection are performed. In light of the limited bandwidth of wireless channels, we develop a level-triggered sampling scheme, where each meter transmits only one-bit each time asynchronously. The proposed multi-meter scheme features substantially low communication overhead, while its performance is close to that of the ideal case where distributed meter readings are perfectly available at the central meter.

Index Terms

Power signal disturbance, autoregressive model, change detection, level-triggered sampling.

I. INTRODUCTION

At present the power quality has become a critical security concern for the emerging power grid system, due to the rapidly growing number of equipments that not only generate but also are sensitive to various disturbances. One crucial task is monitoring the power signal for malicious power quality disturbances that could lead to device damage or even network blackout. In general, power quality disturbances include the voltage disturbance and the current disturbance. They both involve deviations of the actual power

[†]S. Li and X. Wang are with Electrical Engineering Department, Columbia University, New York, NY 10027 (e-mail: {shang,wangx}@ee.columbia.edu).

signal from the nominal sinusoidal waveform with prescribed amplitude and frequency, thus can be treated similarly from the signal processing point of view. In this paper, without loss of generality, we consider the voltage disturbance. The real-time disturbance detection is useful in practice in two ways. On one hand, it enables the power system to promptly respond to the detrimental fluctuations caused by the generator and load operations, capacitor bank switching and abrupt environment change (e.g., lightning strike). On the other hand, it also serves as an abnormal data recording trigger. In power grids, it becomes more and more important to record the power quality data for off-line assessment, in order to help the electricity providers to improve their power supply. However, to avoid the huge volume of data storage, the recorder should only capture the informative data segments corresponding to abnormalities or disturbances and not record data when the power signal is normal. Therefore, to trigger the recording process, timely detection of the disturbance is required.

The monitoring procedure is commonly realized by sampling and analyzing in real-time the voltage waveform, which in practice, is always corrupted by noise. If the noise level is comparable to that of the disturbance, it could lead to frequent false alarms or large decision delay, resulting in poor detection performance. To that end, one key task in power quality monitoring is to detect the voltage disturbance from the observed noisy waveform as soon as possible after its occurrence, so that certain protection measures can be taken and/or the recoding process can be triggered immediately.

A. Background

The existing disturbance detection methods can be roughly categorized into non-model-based and model-based approaches [1, Chapter 7] [2]. In particular, the non-model-based approach directly examines the instantaneous changes from the observed waveform. For example, one method is to use a high-pass filter to capture the high-frequency component induced by the abrupt transition when the disturbance occurs. Intuitively, this method becomes ineffective if the waveform changes smoothly. Moreover, the high-frequency component at the transition point can be buried by the noise. The most widely used method thus far is by monitoring the root mean squared (RMS) sequence, which is computed over a sliding window of length W (usually one cycle of the nominal waveform) as follows:

$$Q(t) = \sqrt{\frac{1}{W} \sum_{k=t-W+1}^t y_k^2}, \quad (1)$$

where y_k is the k th sample of the voltage waveform. A disturbance is detected once the current RMS surpasses or falls below a prescribed threshold. Despite its simplicity, the RMS method is effective in detecting the disturbances associated with magnitude change, e.g., voltage sag, voltage swell. However,

it is shown in [3] that the RMS method can miss some transient disturbances that are associated with spectral changes, which is expected because it is based on the energy level of the waveform and is less sensitive to spectral variation. Other methods detect the instantaneous distortion in the frequency domain, mainly by the wavelet transform (WT) or the short-time Fourier transform (STFT) [2, 4]. These methods are naturally more sensitive to the spectral distortion. In addition, both RMS and STFT (or WT) methods have limited time resolution due to the window over which the RMS and spectrum are evaluated respectively.

In contrast, the model-based framework detects the change of parameters or the large residual between the observed waveform and the nominal model. A typical model is to characterize the nominal waveform and disturbance by the superposition of a number of sinusoidal waveforms. Then classical spectral estimation methods such as the MUSIC, ESPRIT and Kalman filter can be applied to estimate the parameters of these sinusoidal disturbances [1, 5]. The first two methods both require an observation window to estimate the parameters, thus decreasing the time resolution of the detection. All three methods heavily rest on the presumed model, e.g., they require the knowledge of the number of sinusoidal waveforms in the model, which makes them less robust. Alternatively, the autoregressive (AR) model is employed for the voltage disturbance detection in [6], which is capable of capturing a broad range of spectral property, whereas the sinusoidal model can only capture a certain fixed number of frequency components. However, the method in [6] is based on examining the residual during the waveform transition, making it vulnerable to noise.

An effective approach to mitigating the noise is to employ the statistical framework of hypothesis testing. A block-sequential Neyman-Pearson test is used in [7] to analyze the cause of the voltage disturbance [7]. That work treats the off-line disturbance classification rather than the online disturbance detection. In [4], the Kolmogorov-Smirnov (KS) test and the likelihood ratio test are employed to denoise the wavelet transform coefficients. It is a fixed-sample size approach rather than a sequential one, thus is less efficient in terms of time resolution. In [8, 9], under the change detection framework, a sequential online approach based on the weighted CUSUM test was introduced, by examining the different distributions of the observed waveforms before and after the occurrence of the disturbance. However, the disturbance signal is treated as independent over time.

Another noteworthy approach to combat the noise is to employ cooperative meters. In reality, disturbances tend to occur to a group of connected electrical buses at the same time, which brings about the opportunity of detecting the disturbance occurrence in a collaborative fashion. That is, by employing multiple meters at these connected electrical buses, one can draw on the diversity across meters to achieve

better detection performance than using a single meter [10]. In the context of deploying wireless Cyber-network for power system monitoring, meter installation and communication protocols are discussed in [11, 12]. However, specific signal processing techniques of cooperative disturbance detection on top of these physical infrastructures is yet to be investigated.

B. Overview

In this paper, we also formulate the online voltage monitoring as a sequential change detection problem. But compared to [8], we build our framework on the time series model, i.e., AR model, which captures the time correlation of the disturbance, and thus provides more realistic characterization. To tackle the main difficulty that the disturbance signal is typically unknown, we propose a change detector based on the generalized local likelihood ratio (GLLR) test. This approach takes a simple form and is more effective than the existing methods.

Moreover, we consider the scenario where multiple meters are employed for cooperatively monitoring the voltage waveform. As mentioned before, these meters are deployed at various locations in the distribution network, where electrical buses are exposed to the same power quality event¹. Despite the robustness against noise, distributed meters inevitably impose communication challenges. Conventionally, meters transmit their local measurement (quantized with multiple bits) to the central meter, where the cooperative detection is performed. The ideal case is that the distributed meter readings are precisely available to the central meter at every sampling instant, i.e., infinite number of bits for quantization, referred to as the centralized multi-meter detection. However, such a centralized setup induces significant amount of communication overhead. To meet the bandwidth constraints, we develop a decentralized detection scheme where it is assumed that the distributed meters and the central meter are linked by low-rate communication channels. In particular, in the proposed decentralized scheme, each meter performs its own GLLR test; samples its decision statistic using the level-triggered sampling; and transmits its sampled local statistic to the central meter using a single bit. The central meter collects the bits from all linked meters and updates the global decision statistic to make the decision. In Section IV, we provide extensive simulation examples to demonstrate that the proposed decentralized detector performs close to the centralized detector, and outperforms the traditional decentralized approach that is based on uniform-in-time sampling and quantization.

¹Although the power quality event strikes a group of buses altogether, the disturbance signals observed at these buses are not necessarily the same. For example, a line fault event could result in voltage sag at one bus but voltage swell at another bus.

The reminder of this paper is organized as follows. In Section II, we formulate the disturbance detection as a sequential change detection problem based on the AR model and develop a generalized local likelihood ratio (GLLR) test. In Section III, we further develop the decentralized GLLR test based on the level-triggered sampling. Simulation results are provided in Section IV and finally, Section V concludes the paper.

II. SEQUENTIAL CHANGE DETECTION OF DISTURBANCES BASED ON THE AR MODEL

In this section, we formulate the online detection of voltage disturbance as a sequential change detection problem with unknown post-change parameters. Then a generalized local likelihood ratio-based detector is derived.

A. AR Modelling of Disturbances

Without loss of generality, we assume that the disturbance occurs at some unknown time $t = t_0$. That is, before t_0 the nominal voltage waveform is a sinusoid with some nominal magnitude, frequency and phase:

$$f_t^{\text{nominal}} = a_0 \sin(2\pi f_0 t + \phi_0). \quad (2)$$

After t_0 , the disturbance distorts the nominal sinusoid. Since the parameters $\{a_0, f_0, \phi_0\}$ take prescribed values (ϕ_0 is obtained by synchronization), the waveform before t_0 is deterministic. Thus we can subtract the nominal waveform (2) from the measurement to obtain a signal consisting only of noise and disturbance [8, 9]. This preprocessing procedure isolates the disturbance signal, which is comparatively weaker than the nominal waveform [13]. Consequently, the post-preprocessing observed signal before the disturbance occurs consists of measurement noise, which is typically modeled as a white Gaussian process, i.e.,

$$y_t = \nu_t \sim \mathcal{N}(0, \sigma_\nu^2), \quad t < t_0. \quad (3)$$

where y_t is the post-preprocessing meter observation at sampling instant t . After t_0 , the meter observations are comprised of disturbance signal and the measurement noise. To capture the time correlation of the disturbance, we propose to use an autoregressive (AR) model to characterize the disturbance signal, which is popular in analyzing the spectral property of various types of signals, such as speech signals [14] and seismic signals [15], and is applied to model power signal in [6]. The AR model is able to represent a broad spectral range, yielding robust characterization of a variety of potential disturbances; whereas the

sinusoidal model (i.e., modeling the disturbance signal as a sum of sinusoids) in [5] only represents a fixed number of certain frequency components. Moreover, the AR model also requires a small number of parameters and its performance is robust to the model order p [16]. In particular, the post-preprocessing signal after the disturbance occurs can be modeled as

$$\begin{cases} y_t = x_t + \nu_t, \\ x_t = \tilde{\mu} + \sum_{j=1}^p a_j (x_{t-j} - \tilde{\mu}) + w_t, \quad t \geq t_0, \end{cases} \quad (4)$$

where x_t is disturbance signal modeled by an AR process with mean $\tilde{\mu}$, w_t is the driving noise of the AR process and y_t is the meter observations. We further write (4) as

$$y_t = \mu + \sum_{j=1}^p a_j y_{t-j} + u_t, \quad t \geq t_0, \quad (5)$$

where $\mu = \tilde{\mu} (1 - \sum_{j=1}^p a_j)$, and $u_t \triangleq (\nu_t - \sum_{j=1}^p a_j \nu_{t-j} + w_t) \sim \mathcal{N}(0, \sigma_u^2)$ with $\sigma_u^2 = (1 + \sum_{j=1}^p (a_j^t)^2) \sigma_v^2 + \sigma_w^2$ accounts for the excitation of the disturbance and the measurement noise.

Note that the statistical models before and after the disturbance, i.e., (3)-(5), correspond to a standard change detection formulation. The change detection (also termed as the quickest detection [17]) aims at detecting the change point as quickly as possible. It achieves high time-resolution by sequentially observing the measurements in the time domain, and the detection delay is minimized subject to a false alarm constraint. The optimal algorithm is obtained by finding the stopping time T^2 such that

$$\inf_T \sup_{t_0} \mathbb{E}_{t_0} ((T - t_0)^+ | T \geq t_0) \quad \text{subject to} \quad \mathbb{E}_{\infty}(T) \geq \gamma, \quad (6)$$

where \mathbb{E}_{t_0} means the expectation given the change point at t_0 , and \mathbb{E}_{∞} the expectation without any change point. Therefore, the objective function in (6) corresponds to the average detection delay and $\mathbb{E}_{\infty}(T)$ corresponds to the false alarm period, i.e., the time a false alarm appears. Intuitively, (6) aims to minimize the mean detection delay while controlling the period before a false alarm to be longer than γ .

Note that in the change detection formulation, it is assumed that the post-change event is ever-lasting. Although in practice some power quality disturbance is transient in time (i.e., (5) holds for $t_0 < t < t_1$, and after t_1 , the system resumes the normal condition (3)), the expected detection delay is supposed to be much less than the transient duration (such that timely protection measures can be taken before any damage is incurred). Thus we assume t_1 is sufficiently large such that the sequential change detection framework is valid. In Section V, we provide simulation examples to illustrate the performance of the proposed detectors on transient disturbances.

²A stopping time is a random variable, whose value $\{T = t\}$ is determined by the random samples up to t .

Denoting the parameter vector comprised of the AR coefficients and the variance as $\boldsymbol{\theta} = [a_1, a_2, \dots, a_p, \mu, \sigma]^\top$, then the waveform change in (3) and (5) corresponds to the change of the parameter vector at $t = t_0$:

$$\begin{cases} \boldsymbol{\theta} = \boldsymbol{\theta}_0 \triangleq [0, 0, \dots, 0, 0, \sigma_\nu]^\top, & t < t_0, \\ \boldsymbol{\theta} = \boldsymbol{\theta}_1 \triangleq [a_1, a_2, \dots, a_p, \mu, \sigma_u]^\top, & t \geq t_0. \end{cases} \quad (7)$$

Further denoting $\mathbf{y}_j^k \triangleq [y_k, y_{k-1}, \dots, y_j]^\top$, the joint conditional probability density function at time t under these two parameter vectors can be expressed as

$$\begin{aligned} & f_{\boldsymbol{\theta}}(y_t | \mathbf{y}_1^{t-1}) \\ &= \begin{cases} \frac{1}{\sqrt{2\pi}\sigma_\nu} \exp\left(-\frac{\varepsilon_{t,\boldsymbol{\theta}_0}^2}{2\sigma_\nu^2}\right), & \boldsymbol{\theta} = \boldsymbol{\theta}_0 \\ \frac{1}{(\sqrt{2\pi}\sigma_u)} \exp\left(-\frac{\varepsilon_{t,\boldsymbol{\theta}_1}^2}{2\sigma_u^2}\right), & \boldsymbol{\theta} = \boldsymbol{\theta}_1 \end{cases}, \end{aligned} \quad (8)$$

with $\varepsilon_{t,\boldsymbol{\theta}_0} \triangleq y_t$ and $\varepsilon_{t,\boldsymbol{\theta}_1} \triangleq y_t - \mu - \sum_{j=1}^p a_j y_{t-j}$. Then the occurrence of the disturbance can be detected using the following sequential change detection procedure:

$$g_k = \max_{1 \leq j \leq k} S_j^k, \quad T = \inf\{k : g_k \geq h\}, \quad (9)$$

where h is a decision threshold and

$$\begin{aligned} S_j^k &\triangleq \log \frac{f_{\boldsymbol{\theta}_1}(\mathbf{y}_j^k | \mathbf{y}_1^{j-1})}{f_{\boldsymbol{\theta}_0}(\mathbf{y}_j^k | \mathbf{y}_1^{j-1})} = \sum_{i=j}^k \log \frac{f_{\boldsymbol{\theta}_1}(y_i | \mathbf{y}_{i-p}^{i-1})}{f_{\boldsymbol{\theta}_0}(y_i | \mathbf{y}_{i-p}^{i-1})} \\ &= \sum_{i=j}^k \underbrace{\left[\frac{1}{2} \log \frac{\sigma_\nu^2}{\sigma_u^2} - \frac{\varepsilon_{i,\boldsymbol{\theta}_1}^2}{2\sigma_u^2} + \frac{\varepsilon_{i,\boldsymbol{\theta}_0}^2}{2\sigma_\nu^2} \right]}_{s_i}. \end{aligned} \quad (10)$$

Given a target false alarm period γ , the threshold is given by $h \approx \ln(\gamma)$ [16]. Note that at each time k , the test statistic g_k is computed and compared with the threshold h . T is the first time that g_k exceeds h and when the disturbance is declared to occur. If both $\boldsymbol{\theta}_0$ and $\boldsymbol{\theta}_1$ are exactly known, (9) is equivalent to the CUSUM test and the decision statistic g_k can be recursively computed as

$$g_k = (g_{k-1} + s_k)^+, \quad (11)$$

where $(x)^+ \triangleq \max\{x, 0\}$. In essence, (11) forgets the past (i.e., resets) whenever $g_k \leq 0$, which relieves the detector from storing all previous observations in memory compared to that in (9). To that end, (11) can also be represented in the following equivalent form [16, Chapter 2]:

$$N_k = N_{k-1} \mathbb{1}_{\{g_{k-1} > 0\}} + 1, \quad (12)$$

$$g_k = \left(S_{k-N_k+1}^k \right)^+, \quad (13)$$

where N_k is number of observations at time k since the last time of reset at time $k - N_k$ and $\mathbb{1}_{\{\cdot\}}$ is the indicator function.

However, in the disturbance detection problem considered here, the post-change parameter θ_1 is unknown since many types of disturbances may potentially occur. Hence the CUSUM test cannot be directly applied here. One solution is the weighted CUSUM, which requires a prior distribution on the unknown parameters and averages the decision statistic with respect to this presumed prior [16, Chapter 2.4.2]. While the choice of the prior affects the performance substantially, there is no well justified prior available for the various power disturbance signals. Moreover, the computational complexity is usually high due to the multidimensional integral with respect to the prior density. The work [8] adopts this method and assumes that the parameters are independent over time for reasonable computational complexity. Yet another approach is the generalized log-likelihood ratio (GLR) test, where we substitute θ_1 with its maximum-likelihood (ML) estimate in (13):

$$g_k = \left(\sup_{\theta_1} S_{k-N_k+1}^k \right)^+. \quad (14)$$

This is a desirable method when the statistical property of post-change parameters is not known a priori. Still, the maximization in (14) is difficult to solve since $S_{k-N_k+1}^k$ given by (10) is not a convex function of the parameters. Moreover, solving (14) at every time k leads to high computational complexity. In this paper, we apply a generalized local likelihood ratio (GLLR) test to solve our problem, which is elaborated in the next subsection.

B. Generalized Local Likelihood Ratio Test

The GLLR detector is based on the assumption that the parameter change in (7) is small, i.e., $\theta_1 \approx \theta_0$. The corresponding test under this assumption is called the locally optimal test, meaning that the detector is asymptotically optimal as $\theta_1 \rightarrow \theta_0$. Since we have no prior knowledge of the disturbance, assuming that the change is small corresponds to the worst-case scenario that is most difficult to detect. On the other hand, if the disturbance induces significant changes, then it can be easily detected by any simple detection schemes.

The key idea of the local approach is to approximate the decision statistic by a linear expansion around the nominal (pre-change) parameter. Thus we begin by expanding the conditional log-likelihood ratio in

(10) up to the second order, yielding

$$\begin{aligned}
S_j^k &\approx \bar{S}_j^k = \mathbf{r}^\top \left(\sum_{i=j}^k \frac{\partial s_i}{\partial \boldsymbol{\theta}_1} \Big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0} \right) + \frac{1}{2} \mathbf{r}^\top \left(\sum_{i=j}^k \frac{\partial^2 s_i}{\partial \boldsymbol{\theta}_1^2} \Big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0} \right) \mathbf{r} \\
&= \mathbf{r}^\top \left(\sum_{i=j}^k \mathbf{z}_i \right) - \frac{1}{2} \mathbf{r}^\top \left(\sum_{i=j}^k \mathbf{w}_i \right) \mathbf{r},
\end{aligned} \tag{15}$$

where $\mathbf{r} \triangleq \boldsymbol{\theta}_1 - \boldsymbol{\theta}_0$ and

$$\begin{aligned}
\mathbf{z}_i &\triangleq \frac{\partial s_i}{\partial \boldsymbol{\theta}_1} \Big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0} \\
&= \partial \left(-\frac{\varepsilon_{i,\boldsymbol{\theta}_1}^2}{2\sigma_u^2} - \frac{\log 2\pi\sigma_u^2}{2} \right) / \partial \boldsymbol{\theta}_1 \Big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0} \\
&= \left[\begin{array}{c} \frac{1}{\sigma_u^2} \varepsilon_{i,\boldsymbol{\theta}_1} \mathbf{y}_{i-p}^{i-1} \\ \frac{1}{\sigma_u} \left(\frac{\varepsilon_{i,\boldsymbol{\theta}_1}^2}{\sigma_u^2} - 1 \right) \\ \frac{\varepsilon_{i,\boldsymbol{\theta}_1}}{\sigma_u^2} \end{array} \right] \Big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0} = \left[\begin{array}{c} \frac{1}{\sigma_\nu^2} y_i \mathbf{y}_{i-p}^{i-1} \\ \frac{1}{\sigma_\nu} \left(\frac{y_i^2}{\sigma_\nu^2} - 1 \right) \\ \frac{y_i}{\sigma_\nu^2} \end{array} \right],
\end{aligned} \tag{16}$$

$$\begin{aligned}
\mathbf{w}_i &\triangleq -\frac{\partial^2 s_i}{\partial \boldsymbol{\theta}_1^2} \Big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0} \\
&= \frac{1}{\sigma_u^2} \left[\begin{array}{ccc} \mathbf{y}_{i-p}^{i-1} \mathbf{y}_{i-p}^{i-1\top} & \frac{2\varepsilon_{i,\boldsymbol{\theta}_1}}{\sigma_u} \mathbf{y}_{i-p}^{i-1} & \mathbf{y}_{i-p}^{i-1} \\ \frac{2\varepsilon_{i,\boldsymbol{\theta}_1}}{\sigma_u} \mathbf{y}_{i-p}^{i-1\top} & \frac{3\varepsilon_{i,\boldsymbol{\theta}_1}^2}{\sigma_u^2} - 1 & \frac{2\varepsilon_{i,\boldsymbol{\theta}_1}}{\sigma_u} \\ \mathbf{y}_{i-p}^{i-1} & \frac{2\varepsilon_{i,\boldsymbol{\theta}_1}}{\sigma_u} & 1 \end{array} \right] \Big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0} \\
&= \frac{1}{\sigma_\nu^2} \left[\begin{array}{ccc} \mathbf{y}_{i-p}^{i-1} \mathbf{y}_{i-p}^{i-1\top} & \frac{2y_i}{\sigma_\nu} \mathbf{y}_{i-p}^{i-1} & \mathbf{y}_{i-p}^{i-1} \\ \frac{2y_i}{\sigma_\nu} \mathbf{y}_{i-p}^{i-1\top} & \frac{3y_i^2}{\sigma_\nu^2} - 1 & \frac{2y_i}{\sigma_\nu} \\ \mathbf{y}_{i-p}^{i-1} & \frac{2y_i}{\sigma_\nu} & 1 \end{array} \right].
\end{aligned} \tag{17}$$

We still need to decide the change direction \mathbf{r} . Note that (15) is a quadratic function of the change direction \mathbf{r} , which implies that after we replace $S_{k-N_k+1}^k$ in (14) with its second-order approximation $\bar{S}_{k-N_k+1}^k$, i.e., $\sup_{\boldsymbol{\theta}_1} S_{k-N_k+1}^k \approx \sup_{\mathbf{r}} \bar{S}_{k-N_k+1}^k$, then the optimization problem can be solved analytically. We rewrite (15) as

$$\begin{aligned}
\bar{S}_j^k &= \mathbf{r}^\top \left(\sum_{i=j}^k \mathbf{z}_i \right) - \frac{(k-j+1)}{2} \mathbf{r}^\top \left(\sum_{i=j}^k \frac{\mathbf{w}_i}{(k-j+1)} \right) \mathbf{r} \\
&\rightarrow \mathbf{r}^\top \left(\sum_{i=j}^k \mathbf{z}_i \right) - \frac{k-j+1}{2} \mathbf{r}^\top \mathbf{J}(\boldsymbol{\theta}_0) \mathbf{r},
\end{aligned} \tag{18}$$

as $k - j + 1 \rightarrow \infty$, corresponding to the case of large number of samples. This is obtained by considering

$$\sum_{i=j}^k \frac{\mathbf{w}_i}{k - j + 1} \rightarrow \mathbb{E}_{\boldsymbol{\theta}}(\mathbf{w}_i), \quad (19)$$

which can be approximated by the local assumption that $\boldsymbol{\theta}_0 \approx \boldsymbol{\theta}_1$, and hence

$$\mathbb{E}_{\boldsymbol{\theta}}(\mathbf{w}_i) \approx \mathbb{E}_{\boldsymbol{\theta}_0}(\mathbf{w}_i) = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{2}{\sigma_v^2} & 0 \\ \mathbf{0} & 0 & \frac{1}{\sigma_v^2} \end{bmatrix}, \quad (20)$$

where \mathbf{I} is the $p \times p$ identity matrix. We denote the matrix in (20) by $\mathbf{J}(\boldsymbol{\theta}_0)$ because it is, in fact, the Fisher information of the parameter $\boldsymbol{\theta}_0$, i.e.,

$$\mathbf{J}(\boldsymbol{\theta}_0) \triangleq \mathbb{E} \left[- \frac{\partial^2 \log f_{\boldsymbol{\theta}}}{\partial \boldsymbol{\theta}^2} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \right]. \quad (21)$$

We now constrain \mathbf{r} to be on a small ellipse, i.e., $\mathbf{r}^\top \mathbf{J}(\boldsymbol{\theta}_0) \mathbf{r} = b^2$. Based on (18), we can then compute the change direction analytically as follows,

$$\tilde{S}_j^k \triangleq \sup_{\mathbf{r}^\top \mathbf{J}(\boldsymbol{\theta}_0) \mathbf{r} = b^2} \bar{S}_j^k = \sup_{\mathbf{r}^\top \mathbf{J}(\boldsymbol{\theta}_0) \mathbf{r} = b^2} \mathbf{r}^\top \left(\sum_{i=j}^k \mathbf{z}_i \right) - \frac{k - j + 1}{2} \mathbf{r}^\top \mathbf{J}(\boldsymbol{\theta}_0) \mathbf{r}. \quad (22)$$

Decomposing $\mathbf{J}(\boldsymbol{\theta}_0) = \mathbf{B}^\top \mathbf{B}$ with

$$\mathbf{B} \triangleq \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{\sqrt{2}}{\sigma_v} & 0 \\ \mathbf{0} & 0 & \frac{1}{\sigma_v} \end{bmatrix}, \quad (23)$$

it follows that $\mathbf{r}^\top \mathbf{J}(\boldsymbol{\theta}_0) \mathbf{r} = (\mathbf{B}\mathbf{r})^\top (\mathbf{B}\mathbf{r}) = \tilde{\mathbf{r}}^\top \tilde{\mathbf{r}} = b^2$. Then (22) becomes

$$\begin{aligned} \tilde{S}_j^k &= \sup_{\tilde{\mathbf{r}}^\top \tilde{\mathbf{r}} = b^2} \bar{S}_j^k = \sup_{\tilde{\mathbf{r}}^\top \tilde{\mathbf{r}} = b^2} \tilde{\mathbf{r}}^\top (\mathbf{B}^\top)^{-1} \left(\sum_{i=j}^k \mathbf{z}_i \right) - \frac{k - j + 1}{2} b^2 \\ &= \sup_{\tilde{\mathbf{r}}^\top \tilde{\mathbf{r}} = b^2} \tilde{\mathbf{r}}^\top \left(\sum_{i=j}^k \tilde{\mathbf{z}}_i \right) - \frac{k - j + 1}{2} b^2 \\ &= b U_j^k - \frac{k - j + 1}{2} b^2, \end{aligned} \quad (24)$$

with

$$U_j^k \triangleq \left\| \sum_{i=j}^k \tilde{\mathbf{z}}_i \right\| \quad \text{and} \quad \tilde{\mathbf{z}}_i \triangleq (\mathbf{B}^\top)^{-1} \mathbf{z}_i = \begin{bmatrix} \frac{1}{\sigma_v^2} y_i \mathbf{y}_{i-p}^{i-1} \\ \frac{1}{\sqrt{2}} \left(\frac{y_i^2}{\sigma_v^2} - 1 \right) \\ \frac{y_i}{\sigma_v} \end{bmatrix}. \quad (25)$$

Note that the statistic $\tilde{\mathbf{z}}_i$ is a function of the observations and intuitively relates to the disturbance model given by (4). In specific, the term of $y_i \mathbf{y}_{i-p}^{i-1} / \sigma_\nu^2$ accounts for the autocorrelation of meter samples, $(\frac{y_i^2}{\sigma_\nu^2} - 1) / \sqrt{2}$ captures the variance increase of meter samples and y_i / σ_ν monitors the mean shift.

Finally, incorporating the generalized local likelihood ratio statistics (24)-(25) into the sequential change detector (12)-(13), we summarize the GLLR detector for voltage disturbance detection as

$$N_k = N_{k-1} \mathbb{1}_{\{\tilde{g}_{k-1} > 0\}} + 1, \quad (26)$$

$$\tilde{g}_k = \left(\tilde{S}_{k-N_k+1}^k \right)^+, \quad (27)$$

$$\tilde{T} = \inf \{k : \tilde{g}_k \geq h\}. \quad (28)$$

We highlight the following features of the above disturbance detector:

- Benefiting from the ML estimate of the unknown parameters, it can track the disturbance adaptively, thus is robust to a variety of distortions.
- The decision statistic is an analytical function of the observations given by (24)-(25), which is derived based on the assumption of small r , thus is easy to compute. Moreover, similar to the CUSUM test, it has a recursive form and thus can be efficiently implemented.
- It is implemented by sequentially observing the meter readings, thus is expected to outperform the RMS method and the STFT method in terms of detection delay, which rely on a sampling window, within which the RMS and spectrum are evaluated respectively.

In Section IV, we provide extensive experiments based on the realistic disturbance signals to demonstrate the superior performance of the proposed GLLR detector compared to the existing methods.

III. COOPERATIVE MONITORING BASED ON LEVEL-TRIGGERED SAMPLING

In this section, we consider the cooperative detection of disturbances in an area where the buses are exposed to power signal disturbance at the same time. Cooperative detection takes advantage of the fact that a number of electrical buses can be affected at the same by a power quality event, and allows us to achieve faster decision. Particularly, it is implemented by deploying multiple meters across the network that communicate wirelessly with a central meter which is responsible for monitoring the power quality. Consider L meters that are linked wirelessly with a central meter and perform the cooperative disturbance detection. The straightforward scheme is to make the distributed measurements fully available to the central meter by transmitting very finely (infinite-bit) quantized measurements at every sampling instant, i.e., the centralized setup. However, in practice, the wireless links between the distributed meters

and the central meter are characterized by limited bandwidth. Therefore, in designing a practical system, two constraints need to be considered, namely the *rate constraint* (i.e., the distributed meters should communicate with the central meter at a lower rate than the local sampling rate) and the *quantization constraint* (i.e., each meter should transmit a small number of bits every time it communicates with the central meter). In particular, considering the high sampling rate at distributed meters (e.g., for 60Hz AC supply in North America, the sampling rate could be 3.6kHz with 64 samples per cycle) and the large number of quantization bits in order to achieve an acceptable accuracy at the central meter, it is inefficient to inform the central meter of the local observations at every sampling instant. Thus the decentralized detection, where the distributed meters communicate with the central meter in some low-rate fashion, becomes necessary. In this section, we propose a level-triggered sampling scheme which efficiently lowers the communication overhead in terms of both the communication frequency and the number of information bits at each transmission, while preserving the time resolution of the disturbance detection.

We begin with deriving the centralized multi-meter GLLR detection scheme. The pre-change and post-change signal model in the multi-meter setup is written as

$$\begin{cases} y_t^{(\ell)} = \nu_t^{(\ell)} & t < t_0, \\ y_t^{(\ell)} = \mu^{(\ell)} + \sum_{j=1}^p a_j^{(\ell)} y_{t-j}^{(\ell)} + u_t^{(\ell)}, & t \geq t_0, \end{cases} \quad \ell = 1, 2, \dots, L. \quad (29)$$

Here we assume that $\boldsymbol{\theta}^{(\ell)} = [a_1^{(\ell)}, \dots, a_p^{(\ell)}, \mu^{(\ell)}, \sigma_u^{(\ell)}]$ varies with ℓ , because, in general, the distributed meters observe power voltage at different buses, thus the resulting disturbance signals are not necessarily the same. The driving process $u_t^{(\ell)}$ are assumed to be independent across meters. Correspondingly, the log-likelihood ratio function in (10) becomes

$$S_j^k \triangleq \sum_{\ell=1}^L \sum_{i=j}^k \underbrace{\left[\frac{1}{2} \log \frac{\sigma_\nu^2}{\sigma_u^2} - \frac{\varepsilon_{i, \boldsymbol{\theta}_1}^{(\ell)2}}{2\sigma_u^2} + \frac{\varepsilon_{i, \boldsymbol{\theta}_0}^{(\ell)2}}{2\sigma_\nu^2} \right]}_{S_j^{k(\ell)}}, \quad (30)$$

where $\varepsilon_{t, \boldsymbol{\theta}_0}^{(\ell)} \triangleq y_t^{(\ell)}$ and $\varepsilon_{t, \boldsymbol{\theta}_1}^{(\ell)} \triangleq y_t^{(\ell)} - \mu^{(\ell)} - \sum_{j=1}^p a_j^{(\ell)} y_{t-j}^{(\ell)}$. Moreover, the generalized log-likelihood ratio in (14) is evaluated as

$$\tilde{S}_j^k \triangleq \sup_{\boldsymbol{\theta}_1^{(\ell)}, \ell=1, \dots, L} S_j^k = \sum_{\ell=1}^L \left(\sup_{\boldsymbol{\theta}_1^{(\ell)}} S_j^{k(\ell)} \right). \quad (31)$$

Applying the same local approximation as in the last section to $\sup_{\boldsymbol{\theta}_1^{(\ell)}} S_j^{k(\ell)}$, then we further evaluate

(31) as follows:

$$\tilde{S}_j^k = \sum_{\ell=1}^L \underbrace{\left(bU_j^{k(\ell)} - \frac{k-j+1}{2}b^2 \right)}_{\tilde{S}_j^{k,(\ell)}}, \quad (32)$$

where

$$U_j^{k(\ell)} \triangleq \left\| \sum_{i=j}^k \tilde{\mathbf{z}}_i^{(\ell)} \right\|^2, \quad \tilde{\mathbf{z}}_i^{(\ell)} \triangleq \begin{bmatrix} \frac{1}{\sigma_v^2} y_i^{(\ell)} (\mathbf{y}^{(\ell)})_{i-p}^{i-1} \\ \frac{1}{\sqrt{2}} \left(\frac{y_i^{(\ell)^2}{\sigma_v^2} - 1 \right) \\ \frac{y_i^{(\ell)}}{\sigma_v} \end{bmatrix}, \quad (33)$$

and $(\mathbf{y}^{(\ell)})_j^k \triangleq [y_k^{(\ell)}, y_{k-1}^{(\ell)}, \dots, y_j^{(\ell)}]^\top$ denote the observations from sample k to j at sensor ℓ . As a result, the cooperative multi-meter GLLR detector is summarized as

$$N_k = N_{k-1} \mathbb{1}_{\{\tilde{g}_{k-1} > 0\}} + 1, \quad (34)$$

$$\tilde{g}_k = \left(\sum_{\ell=1}^L \tilde{S}_{k-N_k+1}^{k,(\ell)} \right)^+, \quad (35)$$

$$\tilde{T} = \inf \{k : \tilde{g}_k \geq h\}. \quad (36)$$

Comparing (34)-(36) with (26)-(28), we find that the cooperative GLLR detector differs from the single-meter GLLR detector by summing distributed $\tilde{S}_{k-N_k+1}^{k,(\ell)}$ instead of only using that at one meter. As such, the centralized GLLR detector requires the distributed meters to quantize and transmit the local statistic $\tilde{S}_{k-N_k+1}^{k,(\ell)}$ to the central meter at every local sampling instant k .

We next consider the decentralized implementation of (34)-(36). In the decentralized setup, it is important to devise an efficient communication scheme between the distributed meters and the central meter, by which the local statistics are sent to the central meter less frequently and using small number of bits at each transmission. In the following subsections, we propose an efficient decentralized implementation based on the level-triggered communication scheme at the sensors and its associated decision rule at the central meter.

A. Decentralized Detection Based on Level-Triggered Sampling

We first describe the level-triggered sampling strategy, which is essentially a single-bit quantization of the local statistic, where the transmission of the local statistic is only triggered once it hits a certain value, thus is observation-adaptive. Moreover, all meters communicate with the central meter asynchronously, which avoids the use of a global clock for synchronization.

We simplify the notation of the local test statistic at the ℓ th meter $\{\tilde{S}_{k-N_k+1}^{k,(\ell)}\}$ as $\{\tilde{S}_k^{(\ell)}\}$, because N_k is uniquely determined by k , and denote the n th communicating time of the ℓ th meter as k_n^ℓ . Note that at the ℓ th meter, we can decompose the test statistic as

$$\tilde{S}_k^{(\ell)} = \tilde{S}_k^{(\ell)} - \tilde{S}_{k_{n-1}^\ell}^{(\ell)} + \tilde{S}_{k_{n-1}^\ell}^{(\ell)} - \dots - \tilde{S}_{k_1^\ell}^{(\ell)} + \tilde{S}_{k_1^\ell}^{(\ell)} - \tilde{S}_0^{(\ell)}, \quad (37)$$

where $\tilde{S}_0^{(\ell)} = 0$, and k_n^ℓ is recursively defined as

$$k_n^\ell \triangleq \inf \left\{ k > k_{n-1}^\ell : \tilde{S}_k^{(\ell)} - \tilde{S}_{k_{n-1}^\ell}^{(\ell)} \notin (-\underline{\Delta}, \overline{\Delta}) \right\}, \quad k_0^\ell = 0, \quad \tilde{S}_0^{(\ell)} = 0, \quad (38)$$

where $\overline{\Delta}$ and $\underline{\Delta}$ are positive constants, selected to control the frequency of transmission and known to the central meter. According to (38), each meter informs the central meter of its local statistic every time it cumulates to exit the interval $[-\underline{\Delta}, \overline{\Delta}]$. In the ideal case, $\tilde{S}_{k_n^\ell}^{(\ell)} - \tilde{S}_{k_{n-1}^\ell}^{(\ell)}$ hits the boundary exactly in (38), i.e., $\tilde{S}_{k_n^\ell}^{(\ell)} - \tilde{S}_{k_{n-1}^\ell}^{(\ell)} = -\underline{\Delta}$ or $\overline{\Delta}$. Then the local statistic can be delivered by sending only one-bit information of which boundary is hit to the central meter. In particular, the n th one-bit message transmitted by the ℓ th meter is given by

$$x_n^{(\ell)} = \begin{cases} 1, & \text{if } \tilde{S}_{k_n^\ell}^{(\ell)} - \tilde{S}_{k_{n-1}^\ell}^{(\ell)} \geq \overline{\Delta}, \\ -1, & \text{if } \tilde{S}_{k_n^\ell}^{(\ell)} - \tilde{S}_{k_{n-1}^\ell}^{(\ell)} \leq -\underline{\Delta}. \end{cases} \quad (39)$$

In essence, the central meter uniformly samples the local statistic in its value domain instead of in the time domain to lower the transmission frequency. Moreover, the quantization of local statistic is no longer needed, which substantially decreases the amount of data at each transmission. The level-triggered sampling scheme at each meter is summarized as Algorithm 1. Note that the reset signal in the procedure corresponds to the indicator function in the GLLR test (34): recalling that N_k is the number of observations for computing the local statistic, when the global statistic at the central meter $\tilde{S}_k \leq 0$, a reset signal is broadcast to all meters informing them to reset $N_k = 1$; otherwise, with no reset signal, N_k keeps increasing. The above transmission scheme features an inherent data compression and adaptive communication between the local meters and the central meter. Moreover, the one-bit transmission induces significant savings in bandwidth and transmission power.

On the other side, the central meter receives the information bits from each meter asynchronously and updates the global running statistic as follows:

$$\begin{aligned} \tilde{S}_k &= \tilde{S}_{k-1} + \sum_{\ell=1}^L \left(\mathbb{1}_{\{k=k_n^\ell, x_n^{(\ell)}=1\}} \overline{\Delta} - \mathbb{1}_{\{k=k_n^\ell, x_n^{(\ell)}=-1\}} \underline{\Delta} \right) \\ &= \sum_{\ell=1}^L \sum_{n: k_n^\ell < k} \left(\mathbb{1}_{\{x_n^{(\ell)}=1\}} \overline{\Delta} - \mathbb{1}_{\{x_n^{(\ell)}=-1\}} \underline{\Delta} \right), \end{aligned} \quad (40)$$

Algorithm 1 : Level-triggered sampling of the GLLR statistic at the ℓ th meter

```

1: Initialization:  $k \leftarrow 0$ 
2: Reset:  $\lambda \leftarrow 0, N \leftarrow 1$ 
3: while  $\tilde{S}_{k-N+1}^{k,(\ell)} - \lambda \in (-\underline{\Delta}, \overline{\Delta})$  do
4:    $k \leftarrow k + 1$ 
5:   Check the reset signal broadcasted by the central meter:
6:   if present then
7:     go to line 2
8:   else
9:      $N \leftarrow N + 1$ 
10:  end if
11:  Compute  $\tilde{S}_{k-N+1}^{k,(\ell)}$  by (32)-(33)
12: end while
13: Send  $x_k^{(\ell)} = \text{sign}(\tilde{S}_{k-N+1}^{k,(\ell)} - \lambda)$  to the central meter
14:  $\lambda \leftarrow \tilde{S}_{k-N+1}^{k,(\ell)}$ 
15: Check the reset signal broadcast by the central meter:
16: if present then go to line 2
17: else go to line 3.

```

which is essentially the decentralized counterpart of (32). Every time the global statistic is updated at the central meter, it is used to perform the GLLR test given by (34)-(36). There are two decisions to make, i.e., triggering the alarm that a disturbance is detected or continuing to receive information bits from the meters. The procedure at the central meter is summarized as Algorithm 2.

Through the level-triggered sampling scheme, we efficiently recover the decision statistic at the central meter by collecting local statistics from meters. Specifically, compared to the centralized setup where observations are transmitted at every sampling instant with multiple quantization bits, the level-triggered sampling features lower communication frequency (which can be controlled by the parameters $\underline{\Delta}$ and $\overline{\Delta}$) and one-bit representation of each sample.

B. Enhancement

In this subsection, we consider the realistic case where the local statistics do not exactly hit the local thresholds at each level-triggered sampling instant. Under such circumstance, information loss will incur

Algorithm 2 : Operations at the central meter

```

1: Initialization:  $\tilde{S} \leftarrow 0, k \leftarrow 0$ 
2: while  $\tilde{S} < h$  do
3:    $k \leftarrow k + 1$ 
4:   Listen to the meters and receive information bits, say,  $r_1$  “+1”s and  $r_2$  “-1”s
5:    $\tilde{S} \leftarrow \tilde{S} + r_1 \overline{\Delta} - r_2 \underline{\Delta}$ 
6:   if  $r_1 + r_2 \neq 0$  &  $\tilde{S} \leq 0$  then
7:      $\tilde{S} \leftarrow 0$ 
8:     broadcast the reset signal to all meters
9:   end if
10: end while
11: Trigger the disturbance alarm and broadcast the reset signal to all meters

```

due to the overshoot error. For example, in Fig. 1(a), when the increment of the actual statistic $\tilde{S}_k^{(\ell)}$ at the ℓ th meter first exceeds the upper threshold $\overline{\Delta}$ at time k_1^ℓ , an error $\varepsilon_1 \triangleq \tilde{S}_{k_1^\ell}^{(\ell)} - \tilde{S}_0^{(\ell)} - \overline{\Delta}$ is incurred, where $\tilde{S}_0^{(\ell)} = 0$. The next level-triggered sampling occurs at time k_2^ℓ when $\tilde{S}_{k_2^\ell}^{(\ell)} - \tilde{S}_{k_1^\ell}^{(\ell)} > \overline{\Delta}$ and again an error $\varepsilon_2 \triangleq \tilde{S}_{k_2^\ell}^{(\ell)} - \tilde{S}_{k_1^\ell}^{(\ell)} - \overline{\Delta}$ is incurred, ending up with an overall error $\tilde{S}_{k_2^\ell}^{(\ell)} - \hat{S}_{k_2^\ell}^{(\ell)} = \varepsilon_1 + \varepsilon_2$, where $\hat{S}_k^{(\ell)}$ is the transmitted statistic up to time k by meter ℓ given by (42). A main problem with the level-triggered sampling scheme in (38) is that the overshoot errors accumulate over time. In general, using (37)-(38), we can write the actual statistic at the ℓ th meter as

$$\tilde{S}_k^{(\ell)} = (\Delta_n + \varepsilon_n) + (\Delta_{n-1} + \varepsilon_{n-1}) + \cdots + (\Delta_1 + \varepsilon_1), \quad k_n^\ell \leq k < k_{n+1}^\ell, \quad (41)$$

and the corresponding transmitted statistic

$$\hat{S}_k^{(\ell)} = \Delta_n + \Delta_{n-1} + \cdots + \Delta_1, \quad k_n^\ell \leq k < k_{n+1}^\ell, \quad (42)$$

where $\Delta_i \in \{\overline{\Delta}, -\underline{\Delta}\}$ and ε_i is the error incurred at the level-triggered sampling instant k_i^ℓ . Then we have the overall error $\tilde{S}_{k_n^\ell}^{(\ell)} - \hat{S}_{k_n^\ell}^{(\ell)} = \varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_n$ after n level-triggered samplings, resulting in significant distortion on the reconstructed global statistic at the central meter.

This problem is addressed in [18] by introducing extra bits to quantize the overshoot. Here we propose a new simple but effective method to mitigate the overshoot error accumulation problem while preserving the single-bit transmission feature. In particular, we modify the stopping time (38) to

$$k_n^\ell \triangleq \inf \left\{ k > k_{n-1}^\ell : \tilde{S}_k^{(\ell)} - \tilde{S}_{k_{n-1}^\ell}^{(\ell)} + \varepsilon_{n-1} \notin (-\underline{\Delta}, \overline{\Delta}) \right\}. \quad (43)$$

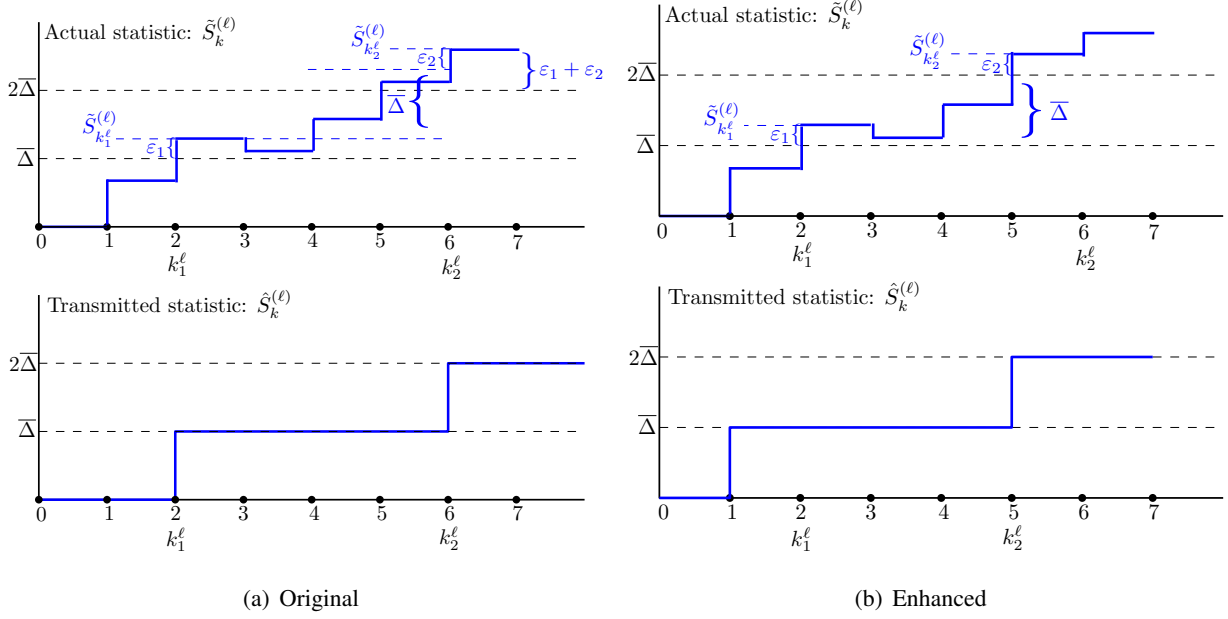


Fig. 1. Illustration of the original and enhanced level-triggered sampling in the presence of overshoot errors.

Under this strategy, at sampling instant k_n^ℓ , the previous overshoot error ε_{n-1} is incorporated in the current sampling. In this way, we have only the current overshoot error present at any time but there is no longer error accumulation. Fig. 1(b) illustrates the modification in (43). The same as in Fig. 1(a), error ε_1 is incurred at the first level-triggered sampling instant k_1^ℓ . However, the next sampling occurs at k_2^ℓ when $\tilde{S}_{k_2^\ell}^{(\ell)} - \tilde{S}_{k_1^\ell}^{(\ell)} + \varepsilon_1 > \bar{\Delta}$. Since ε_1 is included in the transmitted value $\bar{\Delta}$, the overall error after two samplings $\tilde{S}_{k_2^\ell}^{(\ell)} - \hat{S}_{k_2^\ell}^{(\ell)}$ is now only ε_2 . Similarly, ε_2 will be incorporated in the next level-triggered sampling while ε_3 occurs and so on. Simply put, the modification can be interpreted as that, instead of transmitting the increment between the current statistic and the statistic at the last level-triggered sampling instant, we transmit the increment between the current statistic and the total transmitted statistics up to the last sampling instant.

To incorporate (43) into Algorithm 1, the only necessary change is on Line 14, which is replaced by

$$\lambda \leftarrow \lambda + \mathbb{1}_{\{x_n^{(\ell)}=1\}} \bar{\Delta} - \mathbb{1}_{\{x_n^{(\ell)}=-1\}} \underline{\Delta};$$

whereas no change is needed for the operations at the central meter, i.e., Algorithm 2 remains the same.

C. Parameters Tuning

In this subsection, we discuss the selection of the free parameters in our proposed detector, i.e., $\{\underline{\Delta}, \bar{\Delta}, b\}$. The interval $[-\underline{\Delta}, \bar{\Delta}]$ controls the communication frequency between the distributed meters and

the central meter. Denote the communication interval as τ . Then we want to select $[-\underline{\Delta}, \overline{\Delta}]$ such that some target value for $\mathbb{E}_{\theta_0}(\tau)$ is attained, by noting that it is impossible to fix a target value $\mathbb{E}_{\theta_1}(\tau)$ due to the lack of the prior knowledge of the disturbance. We can adjust $[-\underline{\Delta}, \overline{\Delta}]$ by an off-line stochastic simulation to achieve the target $\mathbb{E}_{\theta_0}(\tau)$ based on the statistical distribution of the noise. For the experiments in Section V, we set $\underline{\Delta} = \overline{\Delta}$ and adjust the value to meet the target average transmission interval.

To choose a proper b , first we introduce the detectability condition for the general sequential change detection [16], which states that the two distributions before and after the occurrence of the disturbance are detectable with finite stopping time if and only if $\mathbb{E}_{\theta_0}(\phi_k) < 0 < \mathbb{E}_{\theta_1}(\phi_k)$, where ϕ_k represents the k th sample statistic. This condition implies that the decision statistic $S_1^N = \sum_{k=1}^N \phi_k$ takes different drifting directions between the pre-change and post-change distributions. For the test statistic that cannot be written as the sum of single-sample statistic ϕ_k , the corresponding quantity is also defined in [16] as $\mathbb{E}_{\theta}(\phi_k) \triangleq \lim_{N \rightarrow \infty} \mathbb{E}_{\theta}(S_1^N)/N$, where S_1^N is the statistic with N samples. It is clear that our statistic in (32)-(33) corresponds to the latter definition. Thus at the ℓ th meter we have

$$\begin{aligned} \mathbb{E}_{\theta}(\tilde{\phi}_k^{(\ell)}) &\triangleq \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}_{\theta}(\tilde{S}_1^{N,(\ell)}) \\ &= \lim_{N \rightarrow \infty} \frac{b}{N} \mathbb{E}_{\theta}(U_1^{N,(\ell)}) - \frac{b^2}{2}, \\ \theta &= \theta_0, \theta_1, \ell = 1, 2, \dots, L. \end{aligned} \quad (44)$$

Substituting $U_j^{k,(\ell)}$ with (33), we obtain

$$\begin{aligned} &\lim_{N \rightarrow \infty} \frac{b}{N} \mathbb{E}_{\theta}(U_1^{N,(\ell)}) \\ &= b \mathbb{E}_{\theta} \left(\lim_{N \rightarrow \infty} \left\{ \sum_{m=1}^p \left(\frac{\sum_{i=1}^N y_i^{(\ell)} y_{i-m}^{(\ell)}}{N \sigma_{\nu}^2} \right)^2 + \frac{1}{2} \left(\frac{\sum_{i=1}^N (y_i^{(\ell)})^2}{N \sigma_{\nu}^2} - 1 \right)^2 + \left(\frac{\sum_{i=1}^N y_i^{(\ell)}}{N \sigma_{\nu}} \right)^2 \right\}^{1/2} \right) \\ &= b \rho^{(\ell)}, \end{aligned} \quad (45)$$

where

$$\rho^{(\ell)} \triangleq \left\{ \sum_{m=1}^p \left(\frac{R^{\ell}(m) + (\tilde{\mu}^{(\ell)})^2}{\sigma_{\nu}^2} \right)^2 + \left(\frac{R^{\ell}(0) + (\tilde{\mu}^{(\ell)})^2 - \sigma_{\nu}^2}{\sqrt{2} \sigma_{\nu}^2} \right)^2 + (\tilde{\mu}^{(\ell)})^2 / \sigma_{\nu}^2 \right\}^{1/2}, \quad (46)$$

$R^{\ell}(m)$, $m = 0, 1, \dots, p$ and $\tilde{\mu}^{(\ell)}$ are the autocovariance and the mean shift of the observations at meter ℓ respectively.

We choose b such that the detectability condition holds, i.e., $\mathbb{E}_{\theta_0}(\tilde{\phi}_k^{(\ell)}) < 0 < \mathbb{E}_{\theta_1}(\tilde{\phi}_k^{(\ell)})$, $\ell = 1, 2, \dots, L$. Specifically, before the disturbance occurs, only the white Gaussian noise exists, hence

$\tilde{\mu}^{(\ell)} = 0, R^\ell(0) = \sigma_\nu^2, R^\ell(1) = \dots = R^\ell(p) = 0$, thus we have

$$\mathbb{E}_{\theta_0}(\tilde{\phi}_k^{(\ell)}) = -\frac{b^2}{2} < 0, \quad \ell = 1, 2, \dots, L, \quad (47)$$

for any positive value of b . After the disturbance occurs, we have

$$\mathbb{E}_{\theta_1}(\tilde{\phi}_k^{(\ell)}) = b\rho^{(\ell)} - \frac{b^2}{2} > 0, \quad \ell = 1, 2, \dots, L, \quad (48)$$

for $0 < b < 2 \min \{\rho^{(\ell)}, \ell = 1, \dots, L\}$. Furthermore, it is known that, as the detection threshold $h \rightarrow \infty$, the following approximation holds [16]:

$$\mathbb{E}_{\theta_1}(T) \approx \frac{\ln \gamma}{\sum_{\ell=1}^L \mathbb{E}_{\theta_1}(\tilde{\phi}_k^{(\ell)})} \leq \frac{\ln \gamma}{b \sum_{\ell=1}^L \rho^{(\ell)} - b^2 L/2}, \quad (49)$$

where T denotes the detection delay and γ is the false alarm period. Therefore, given the false alarm period, the mean detection delay is minimized when $\sum_{\ell=1}^L \mathbb{E}_{\theta_1}(\tilde{\phi}_k^{(\ell)})$ attains its maximum at $b = \frac{1}{L} \sum_{\ell=1}^L \rho^{(\ell)}$. In practice, the practitioners can decide the minimum $\rho_{\min}^{(\ell)}$ at each meter and select $b = \frac{1}{L} \sum_{\ell=1}^L \rho_{\min}^{(\ell)}$ such that the detection performance for worst-case disturbance is optimized.

IV. SIMULATION RESULTS

In this section, we first apply the proposed detector on some typical power quality disturbances to demonstrate that it promptly detects the occurrence of these disturbances. In specific, we compare the proposed detector with the widely used RMS method, the STFT method and the weighted CUSUM method in [8]. Then we examine the performance of the proposed cooperative detection scheme as well as its decentralized implementation based on the level-triggered sampling.

In our experiment, the disturbance signals are obtained by constructing simulation systems using the popular Matlab toolbox SimPowerSystems [19]. We mainly consider the disturbance of voltage sag induced by a distribution line fault (simulated by constructing the network in Fig. 2 according to [20]) and the transient disturbance induced by the capacitor bank switching (simulated by constructing the network in Fig. 3 according to [21]). In particular, Fig. 2 corresponds to a simplified distribution network where three-phase power supply is transmitted and distributed to “load 1” and “load 2”. Three meters are deployed to monitor this distribution network. In Fig. 3, the “capacitor 1” and “capacitor 2” constitute the capacitor bank which can be switched on and off to adjust the power factor. Throughout the experiment, The nominal voltage is a sinusoidal waveform with $f_0 = 60\text{Hz}$ and unit magnitude. The sampling rates at all meters are set as the standard 64 samples per cycle, i.e., sampling frequency $f_s = 60 \times 64 \text{ Hz}$. The GLLR detector is applied with a first-order AR model (i.e., $p = 1$) and the parameter is set as $b = 0.5$.

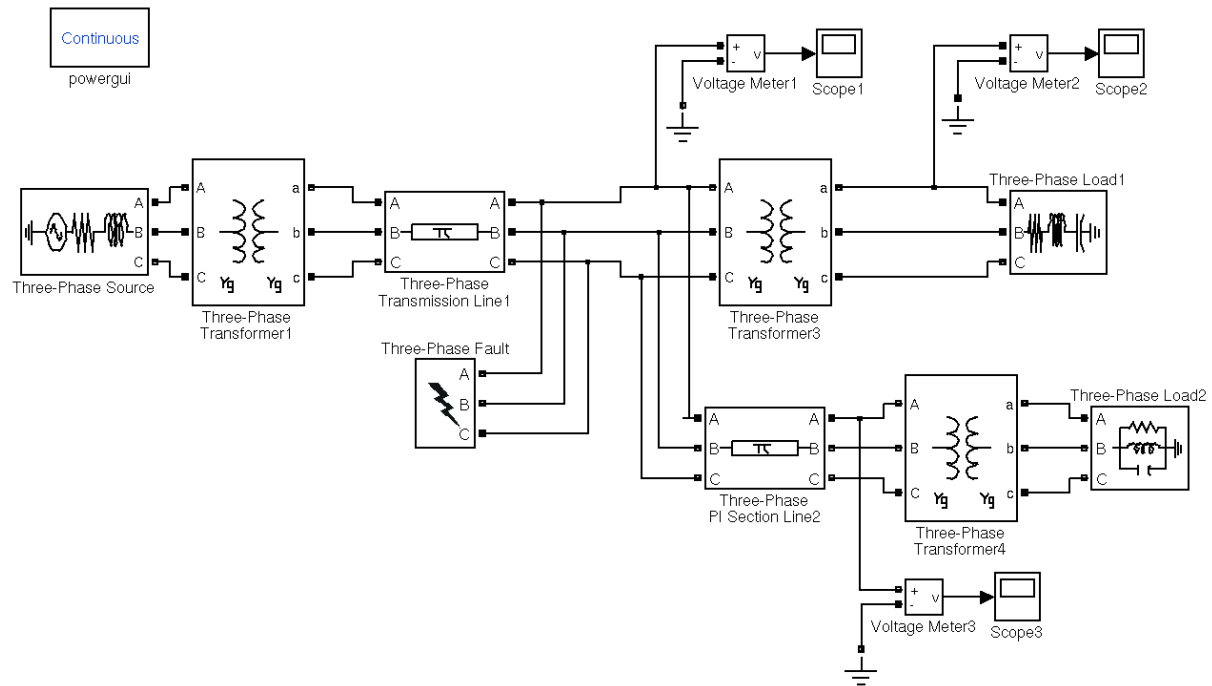


Fig. 2. The simulation system for fault-induced power sag disturbance.

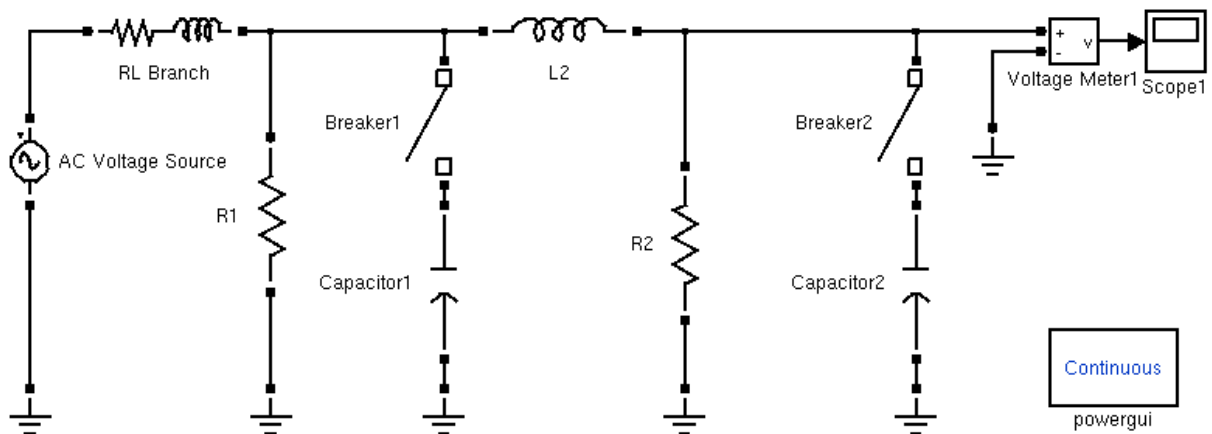


Fig. 3. The simulation system for capacitor-switching-induced transient power disturbance.

A. Comparison with Existing Methods (Single Meter)

By focusing on the single-meter detection, we compare the proposed GLLR detector with the widely adopted methods, namely the RMS method, the STFT method and the weighted CUSUM test. Fig. 4 illustrates the power sag disturbance incurred by the “Phase A line” fault at $t = 0.0869s$ (the occurrence is marked with dashed blue line). Fig. 5 illustrates the power transient distortions incurred by the closing “capacitor 1” in Fig. 3 at $t = 0.105s$. They correspond to the voltage waveform at “Meter 1” in Fig. 2 and Fig. 3 respectively. Both the RMS method and the STFT method are implemented with a one-cycle window that slides point by point, achieving the best possible time resolution. Thus the STFT method performs the 64-point FFT within each window. One can chose a larger window for higher resolution in the frequency domain at the price of lower resolution in the time domain, which is less desirable when quick detection is considered. The weighted CUSUM is implemented using the Gaussian prior as proposed in [8].

Figs. 6-7 show the decision statistics of the proposed GLLR detector, weighted CUSUM (W-CUSUM), RMS detector and STFT detector. For both types of disturbances, the decision statistic of GLLR exhibits abrupt changes on the occurrence of the disturbance. In contrast, the decision statistic of W-CUSUM slowly increases after the occurrence of power sag and fails to detect the transient disturbance. Moreover, in both cases, the STFT detector vaguely shows the presence of new frequency component upon the occurrence of disturbance, and the RMS detector does not indicate the occurrence clearly.

To perform a rigorous comparison, the mean detection delay versus the false alarm period is further examined based on the power sag disturbance shown in Fig. 8. It is seen that the GLLR outperforms the other methods by yielding much shorter mean delay with the same false alarm period. That is, the GLLR detector reacts much quicker to the occurrence of disturbance signal.

B. Cooperative Detector (Centralized and Decentralized)

We next incorporate more meters in the network and examine the performance of cooperative detection. Focusing on the power sag event, the disturbance signals observed at Meters 1-3 are illustrated in Fig. 9. We see that the disturbance signals induced by the same event occur at the same time to multiple buses but vary from each other in terms of the waveform. In Fig. 10, the decision statistics of the single-meter detector (i.e., S-GLLR), the centralized cooperative detector (i.e., C-GLLR), the decentralized detector based on level-triggered sampling (LTS-GLLR) and the enhanced LTS-GLLR (eLTS-GLLR) are plotted. First, the cooperative detector exhibits steeper increase of the decision statistic compared to the single-meter detector, implying a more prompt reaction to disturbance signals. In the mean time, the global

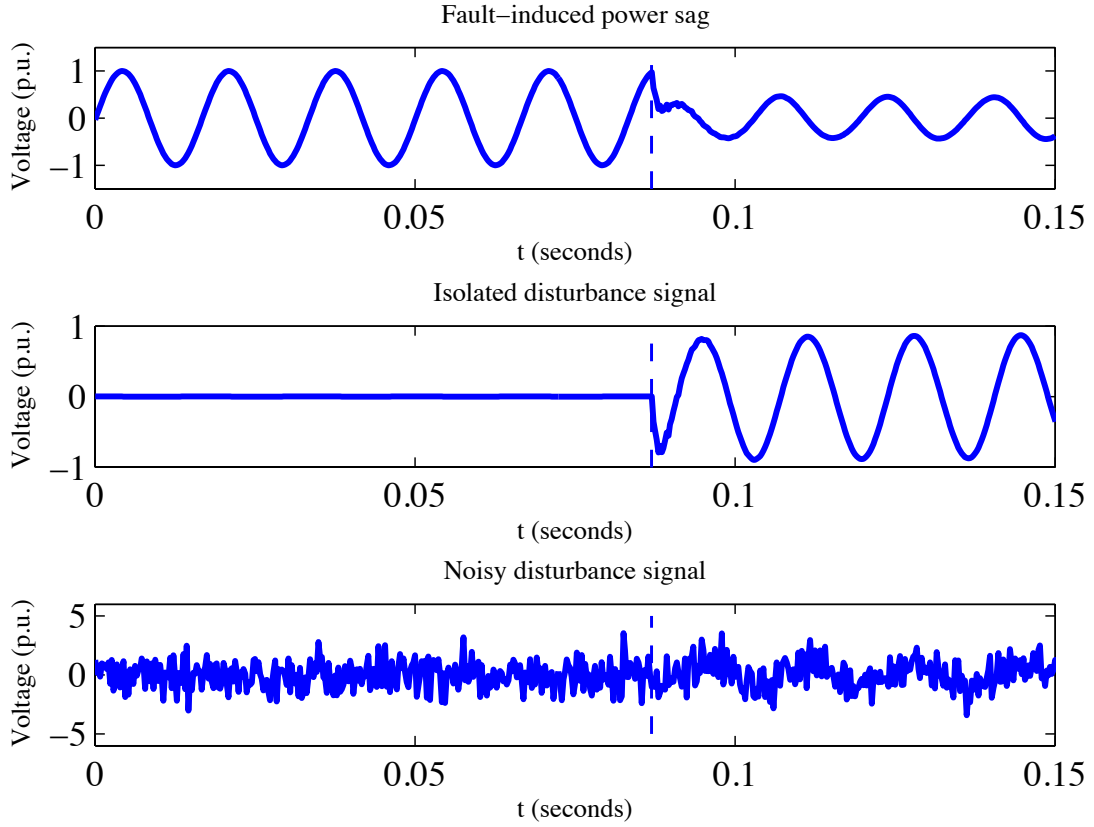


Fig. 4. Original voltage waveform with power sag, isolated disturbance signal and the corresponding noisy measurements.

decision statistics of LTS-based decentralized detectors are updated with a much lower frequency than the centralized one. In particular, the original LTS-GLLR detector clearly diverges from the centralized one due to overshoot accumulation over time, while the enhanced decentralized detector matches closely with the centralized detector. Next we examine the cooperative detectors in terms of detection delay versus the false alarm period. The local thresholds for the level-triggered sampling is chosen as $[-\underline{\Delta}, \overline{\Delta}] = [-1.6, 1.6]$, under which, at each distributed meter, we have $\mathbb{E}_{\theta_0}(\tau) = 14$ samples under normal condition, and $\mathbb{E}_{\theta_1}(\tau) = 4$ samples after the occurrence of disturbance. Compared with the single-meter case, it is seen that cooperative detection with three meters substantially improves the performance in terms of achieving a shorter detection delay. Notably, the proposed LTS-based decentralized detector only exhibits a minor increase of detection delay compared to the centralized detector. As expected, the improvement of eLTS-GLLR over the original LTS-GLLR becomes more significant as the detection delay grows and

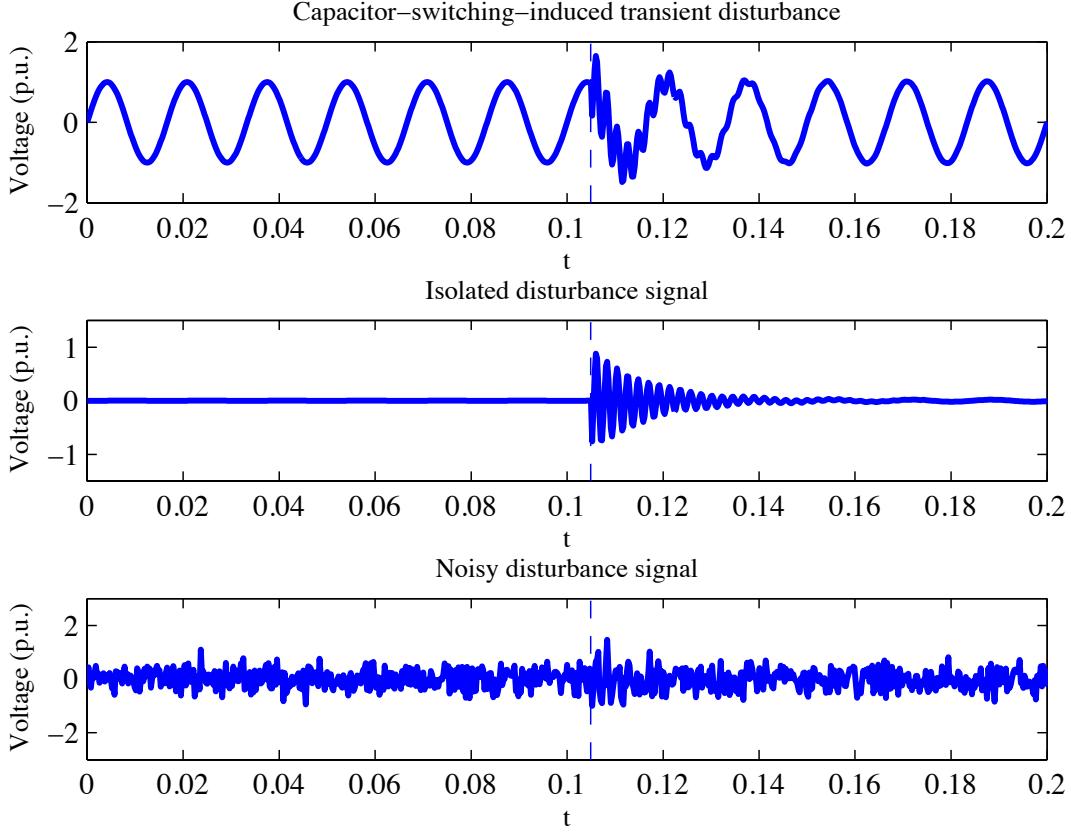


Fig. 5. Original voltage waveform with transient distortion, isolated disturbance signal and the corresponding noisy measurements.

overshoot errors accumulate.

In Fig. 11, we also demonstrate the power of the level-triggered sampling by comparing the proposed decentralized detector with a simple decentralized detector, where each local meter computes its local statistic and transmits it to the central meter every $\tau > 1$ sampling instants (also termed as uniform decentralized detector, which we refer to as U-GLLR in the experiment). Note that when $\tau = 1$, this scheme becomes the centralized detector. Here we set $\tau = 14$ for the simple decentralized detector to match that of the eLTS-GLLR under normal condition. That is, the simple decentralized detector transmits equally frequently as eLTS-GLLR under the normal condition. However, due to the lack of adaptiveness, the time resolution of U-GLLR is limited by τ even in the presence of disturbance signals. Moreover, we assume that in the simple decentralized detector, each local meter transmits the exact value of its local

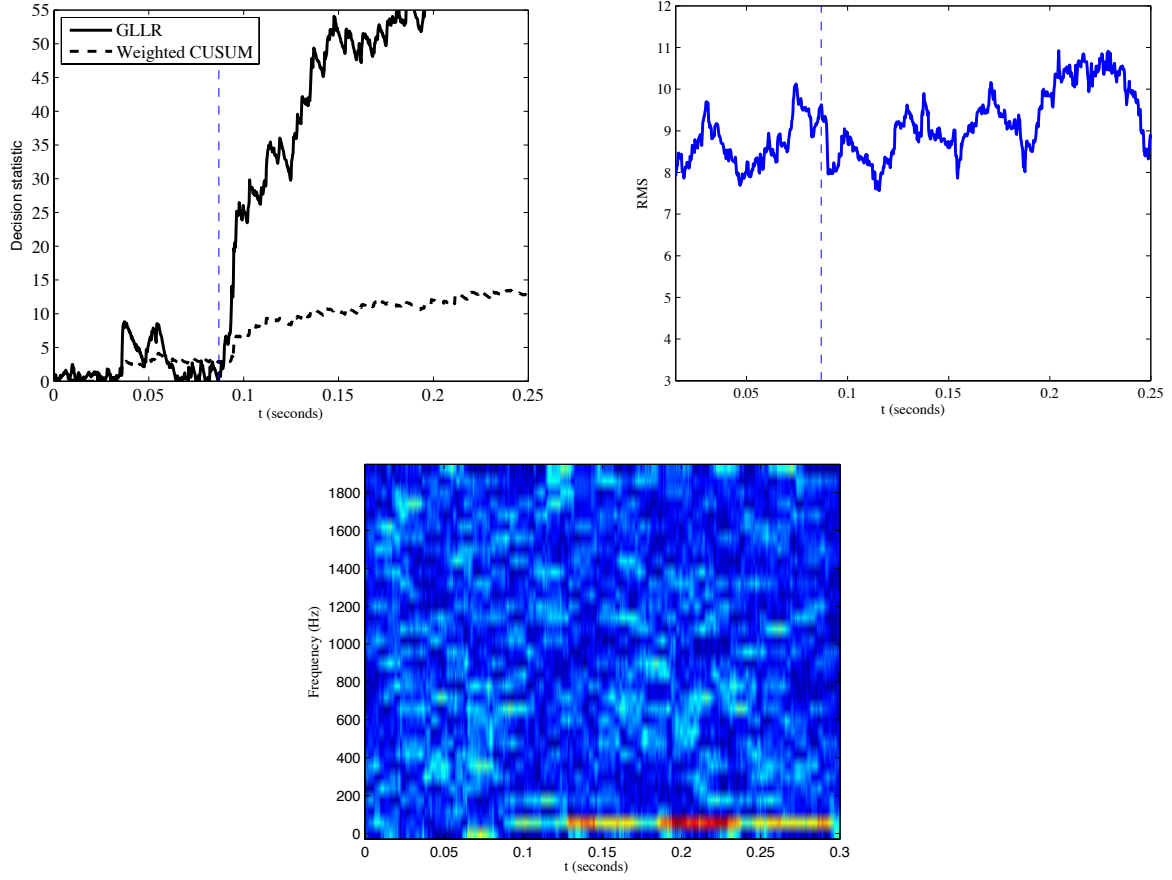


Fig. 6. Detection of power sag disturbance using the GLLR, weighted CUSUM, RMS and STFT methods.

statistic which corresponds to infinite number of bits for each transmission; whereas in the decentralized detectors based on level-triggered sampling, only one bit is sent at each transmission. Remarkably, it is seen in Fig. 11 that even with only one-bit transmission, level-triggered sampling still outperforms the traditional uniform-in-time sampling that transmits infinite number of bits.

Finally, Fig. 12(a)-(b) depict the performances of the centralized and decentralized detectors as the number of meters grows. Again, we consider the power sag event. The communication rate of the simple decentralized detector is fixed at $\tau = 14$, and the average communication rate of proposed decentralized detector is controlled to be $\mathbb{E}_{\theta_0}(\tau) = 14$ and $\mathbb{E}_{\theta_1}(\tau) = 4$. The false alarm period is set as $\gamma = 2 \times 10^3$. It is seen that the detection performance is significantly improved (i.e., smaller mean delay) as the number of meters grows for all methods, implying the benefit of cooperative detection. As the noise level increases, more distributed meters are required to achieve the same detection performance. Note that, due to the

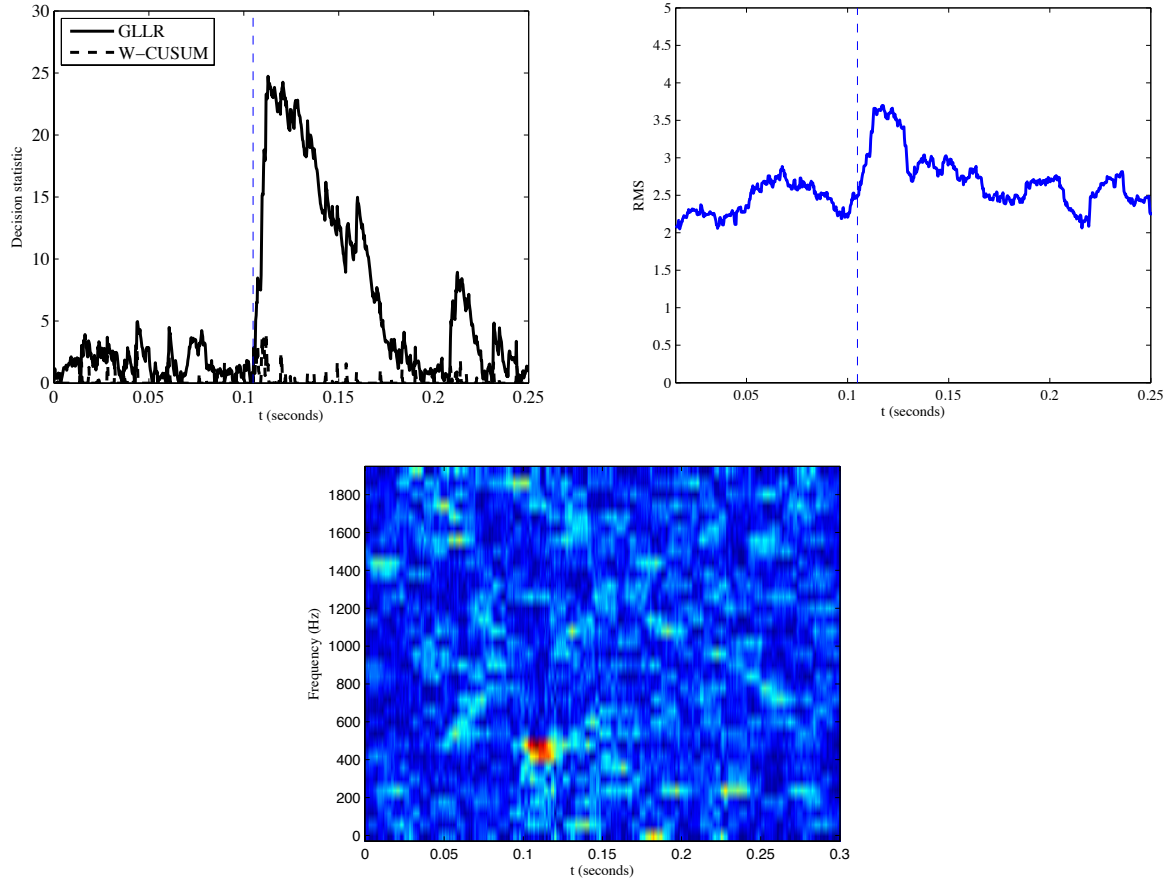


Fig. 7. Detection of transient disturbance using the GLLR, weighted CUSUM, RMS and STFT methods.

lack of adaptiveness, the detection delay of U-GLLR is saturated at mean delay of 14 samples regardless of increasing number of meters. In contrast, eLTS-GLLR consistently outperforms other decentralized detectors.

V. CONCLUSIONS

We have developed a cooperative sequential change detection framework for online power quality monitoring. Specifically, local meters observe the voltage signal independently and communicate wirelessly with a central meter to detect the disturbance. The goal is to achieve the quickest detection under a certain false alarm constraint. First, based on the AR modeling of the disturbance and the sequential change detection framework, we have proposed a sequential GLLR test that does not require the knowledge of the model parameters. Unlike the conventional RMS or STFT method, the proposed

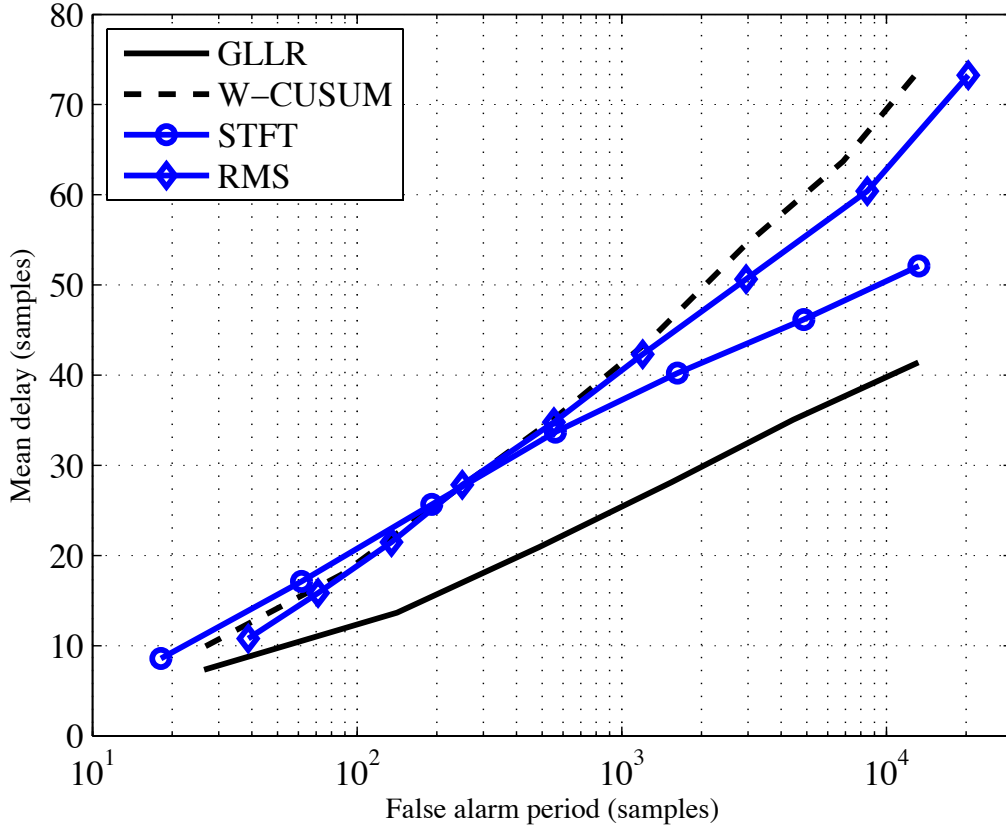


Fig. 8. The detection delay versus the false alarm period for the GLLR detector, weighted CUSUM detector, STFT method and RMS method.

technique exploits the statistical distributions of the observed waveform before and after the occurrence of disturbance, thus provides superior performance, especially in the noisy environment. We have also developed the decentralized version of the GLLR detector, which is specifically tailored toward the low-bandwidth requirement imposed by the wireless transmissions between the distributed meters and the central meter. This is achieved by a novel level-triggered sampling scheme that features single-bit information transmission. Finally we have provided extensive simulation results to demonstrate the superior performance of the proposed centralized and decentralized cooperative detectors over the existing methods.

REFERENCES

- [1] M. H. J. Bollen and I. Y. H. Gu, *Signal processing of power quality disturbances*. Hoboken, NJ: Wiley-IEEE Press, 2006.

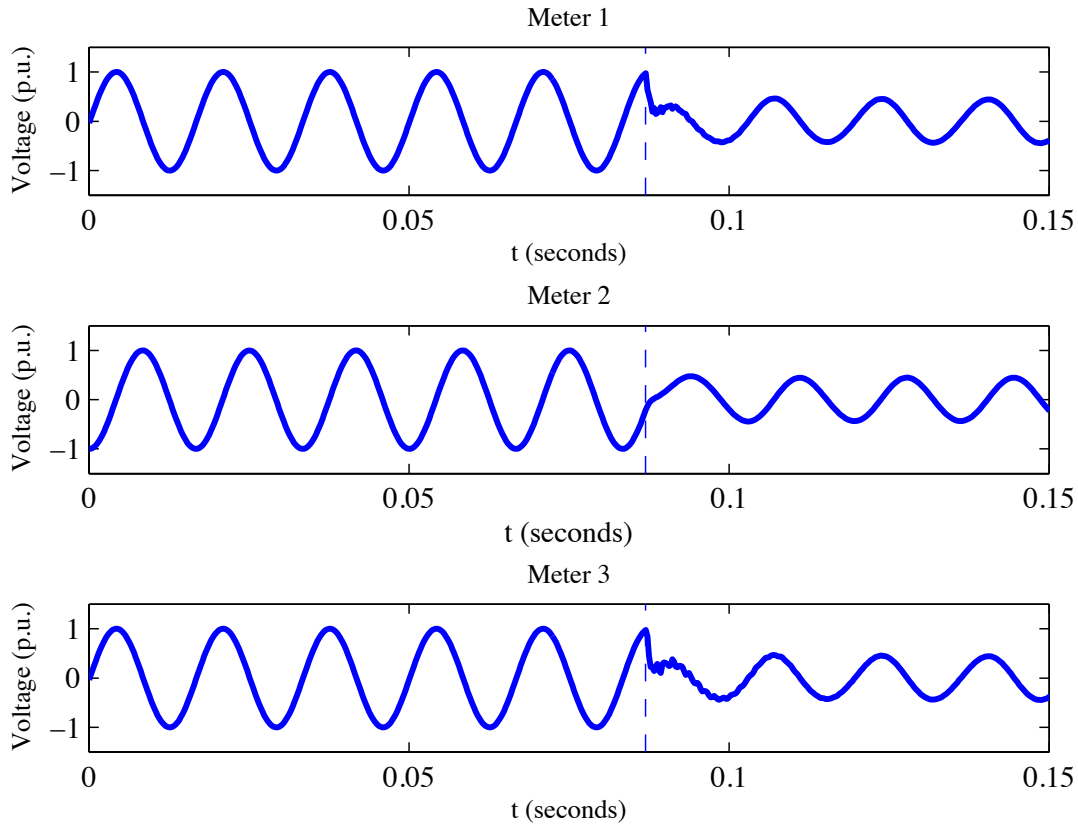


Fig. 9. Original voltage waveform with fault-induced power sag disturbance at Meters 1-3.

- [2] M. H. J. Bollen, I. Y. Gu, and *et al.*, “Bridging the gap between signal and power: assessing power system quality using signal processing techniques,” *IEEE Signal Process. Mag.*, vol. 26, pp. 12–31, July 2009.
- [3] A. M. Gargoom, N. Ertugrul, and W. L. Soong, “A comparative study on effective signal processing tools for power quality monitoring,” in *Proc. 11th Euro. Conf. Power Eletronics and Applications*, Dresden, Germany, Sep. 2005.
- [4] U. D. Dwivedi and S. N. singh, “Denoising techniques with change-point approach for wavelet-based power-quality monitoring,” *IEEE Trans. Power Del.*, vol. 24, no. 3, pp. 1719–1727, July 2009.
- [5] C. J. Dafis, C. O. Nwankpa, and A. Petropulu, “Analysis of power system transient disturbances using an ESPRIT-based method,” in *Proc. IEEE Power Engineering Society Summer Meeting*, Seattle, WA, July 2000.
- [6] I. Y. H. Gu, M. H. J. Bollen, and E. Styvaktakis, “The use of time-varying AR model for the characterization of voltage distrubances,” in *Proc. IEEE Power Engineering Society Winter Meeting*, vol. 4, Jan. 2000, pp. 2943–2948.
- [7] I. Y. H. Gu, N. Ernberg, E. Styvaktakis, and M. H. J. Bollen, “A statistical-based sequential method for fast online detection of fault-induced voltage dips,” *IEEE Trans. Power Del.*, vol. 19, no. 2, pp. 497–504, April 2004.
- [8] X. He, M.-O. Pun, C.-C. Kuo, and Y. Zhao, “A change-point detection approach to power quality monitoring in smart grids,” in *Proc. IEEE International Workshop on Smart Grid Communication*, Cape Town, South Africa, May 2010, pp.

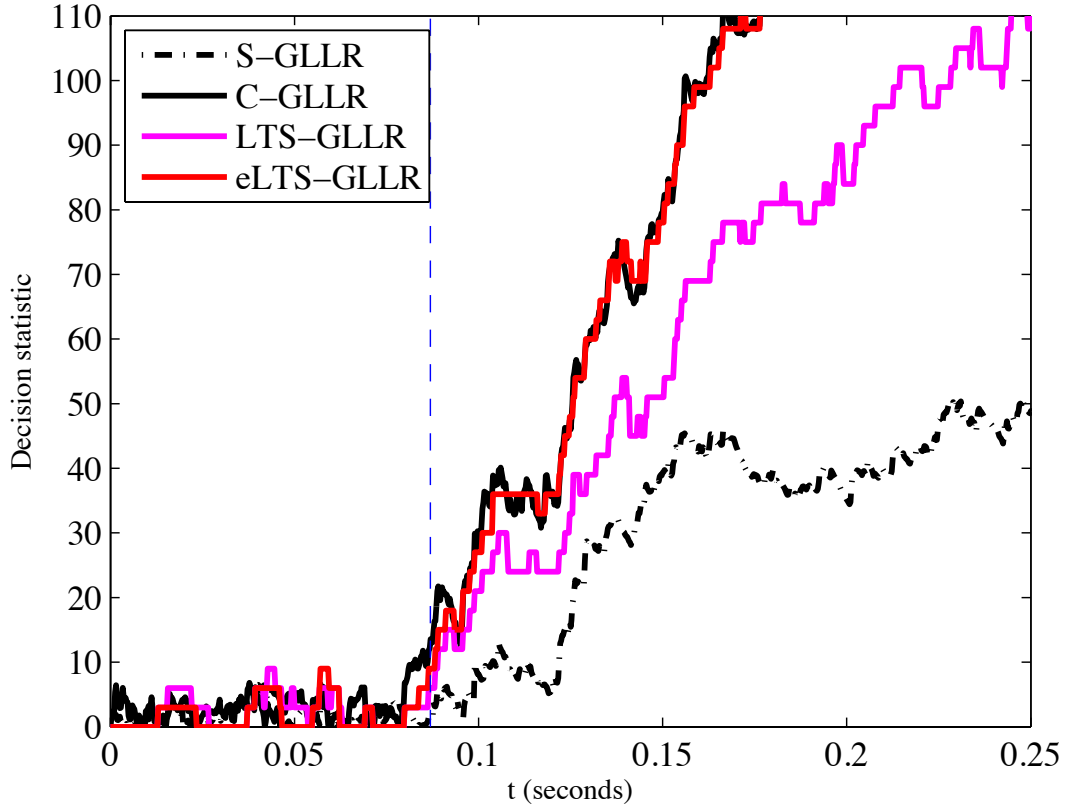


Fig. 10. Decision statistics of the single-meter detector and the cooperative detectors based on level-triggered sampling.

1–5.

- [9] X. He, M.-O. Pun, and C.-C. Kuo, “Quickest detection of unknown power quality events for smart grids,” in *Proc. Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC)*, Hollywood, CA, Dec. 2012, pp. 1–4.
- [10] S. Li and X. Wang, “Monitoring disturbances in smart grids using distributed sequential change detection,” in *Proc. IEEE 5th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, vol. 4, Dec. 2013, pp. 432–435.
- [11] J. Höglund, D. Ilic, S. Karnouskos, R. Sauter, and P. Goncalves Da Silva, “Using a 6LoWPAN smart meter mesh network for event-driven monitoring of power quality,” in *Proc. 3rd IEEE International Conference on Smart Grid Communications (SmartGridComm)*, Tainan City, Taiwan, 5–8 Nov. 2012.
- [12] D. Ilic, S. Karnouskos, and P. Goncalves Da Silva, “Sensing in power distribution networks via large numbers of smart meters,” in *Proc. 3rd IEEE PES Innovative Smart Grid Technologies (ISGT) Europe*, Berlin, Germany, 14–17 Oct. 2012.
- [13] Y.-J. Shin, E. J. Powers, M. Grady, and A. Arapostathis, “Power quality indices for transient disturbances,” *IEEE Trans. Power Del.*, vol. 21, no. 1, pp. 253–261, Jan. 2006.
- [14] R. Andre-Obrecht, “A new statistical approach for the automatic segmentation of continuous speech signals,” *IEEE Trans.*

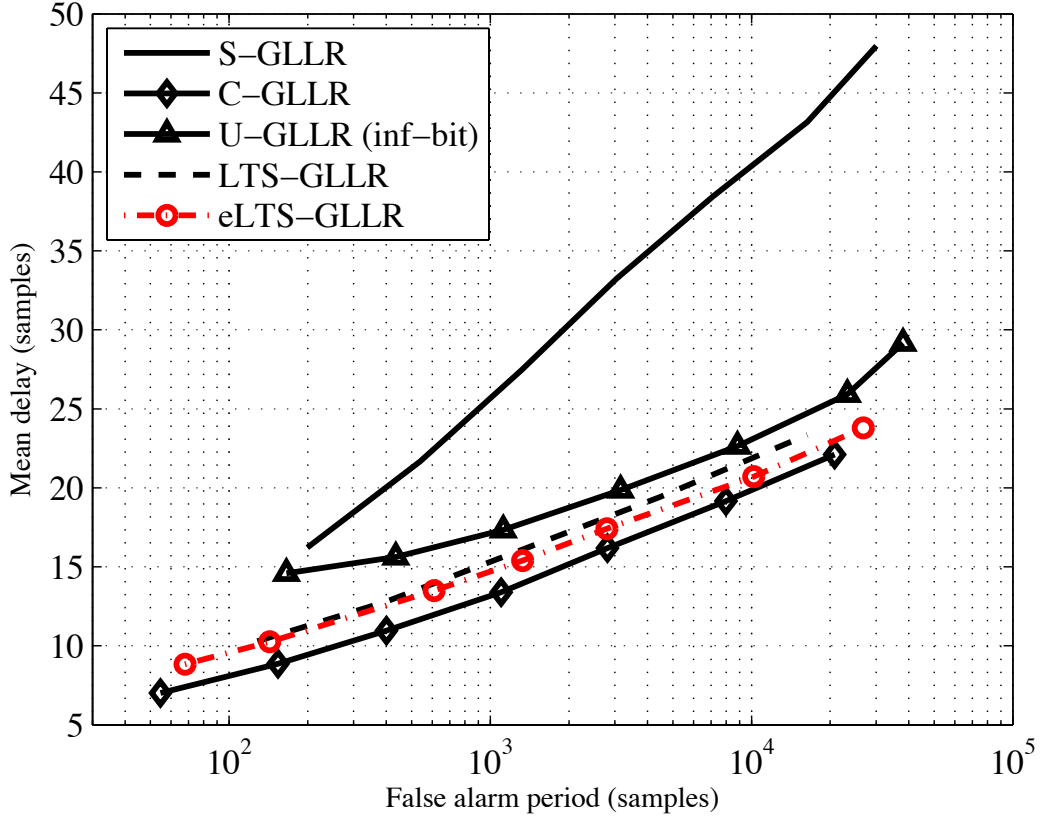
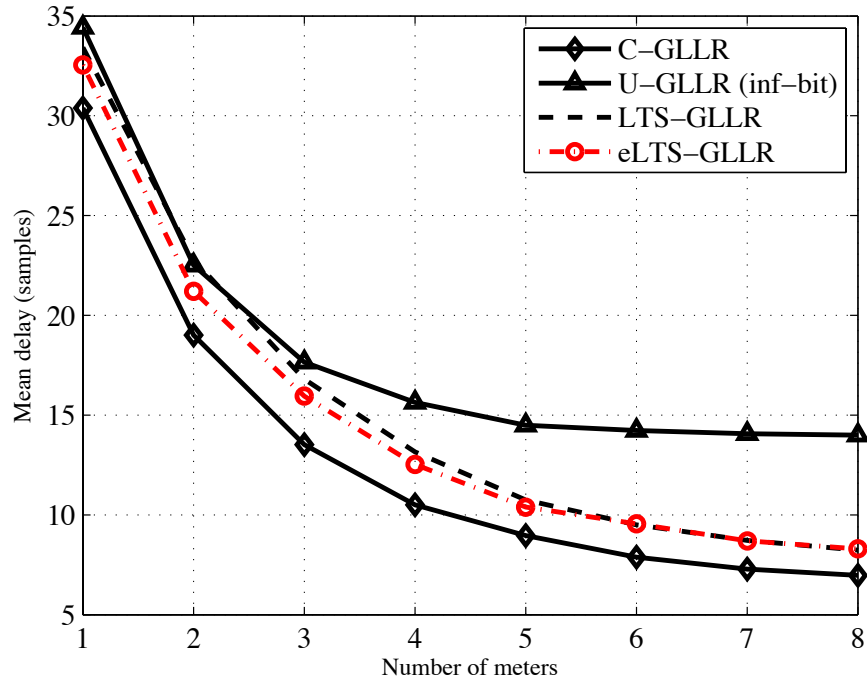
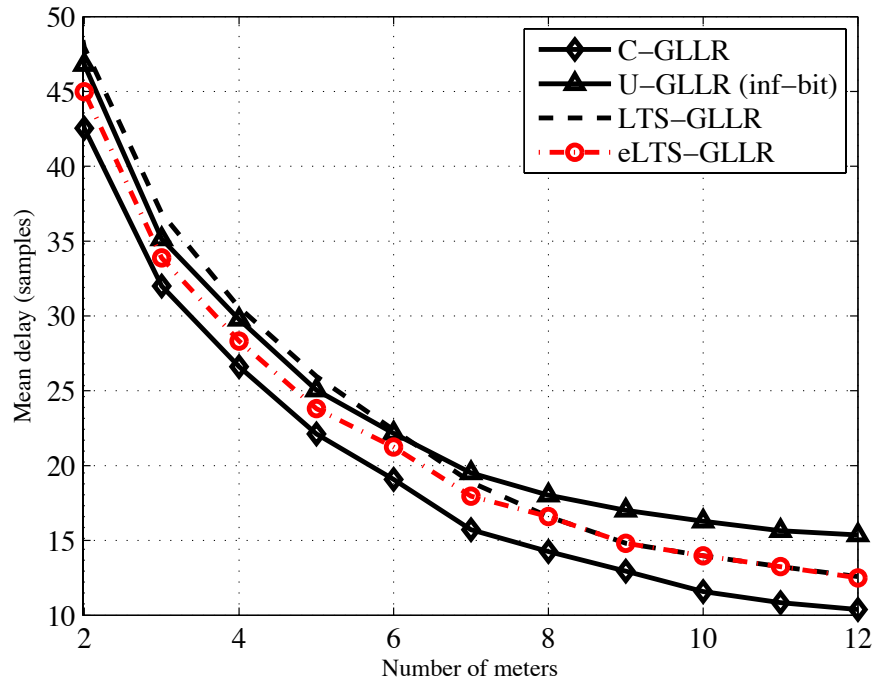


Fig. 11. The detection delay versus the false alarm period for the single-meter detection and the cooperative detection (centralized and decentralized detectors).

Acoust., Speech, Signal Process., vol. 36, no. 1, pp. 29–40, Jan. 1988.

- [15] D. Tjøstheim, “Autoregressive representation of seismic P-wave signals with an application to the problem of short-period discriminants,” *Geophysical Journal of Royal Astronomical Society*, vol. 43, pp. 269–291, 1975.
- [16] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*. Prentice Hall, 1993.
- [17] H. V. Poor and O. Hadjiladis, *Quickest Detection*. Cambridge: Cambridge University Press, 2009.
- [18] Y. Yilmaz, G. Moustakides, and X. Wang, “Cooperative sequential spectrum sensing based on level-triggered sampling,” *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4509–4524, Sep. 2012.
- [19] “SimPowerSystems: For use with Simulink,” [Online]. Available: http://www.mathworks.com/help/releases/R13sp2/pdf_doc/powersys/powersys.pdf.
- [20] R. H. G. Tan and V. K. Ramachandramurthy, “Simulation of power quality events using Simulink model,” in *IEEE 7th International Power Engineering and Optimization Conference (PEOCO2013)*, Langkawi, Malaysia, Jun. 2013.
- [21] G. Gopakumar, H. Yan, B. A. Mork, and K. K. Mustaphi, “Shunt capacitor bank switching transients: A tutorial and case study,” [Online]. Available: http://www.ece.mtu.edu/faculty/bamork/EE5223_S11/mip99.pdf.

(a) $\sigma_v^2 = 0.5$ (b) $\sigma_v^2 = 1$ Fig. 12. The mean detection delay versus increasing number of meters given $\gamma = 2000$.